

# Chameleon: Plug-and-Play Compositional Reasoning with Large Language Models

Year: 2023 | Citations: 403 | Authors: Pan Lu, Baolin Peng, Hao Cheng, Michel Galley, Kai-Wei Chang

---

## Abstract

Large language models (LLMs) have achieved remarkable progress in solving various natural language processing tasks due to emergent reasoning abilities. However, LLMs have inherent limitations as they are incapable of accessing up-to-date information (stored on the Web or in task-specific knowledge bases), using external tools, and performing precise mathematical and logical reasoning. In this paper, we present Chameleon, an AI system that mitigates these limitations by augmenting LLMs with plug-and-play modules for compositional reasoning. Chameleon synthesizes programs by composing various tools (e.g., LLMs, off-the-shelf vision models, web search engines, Python functions, and heuristic-based modules) for accomplishing complex reasoning tasks. At the heart of Chameleon is an LLM-based planner that assembles a sequence of tools to execute to generate the final response. We showcase the effectiveness of Chameleon on two multi-modal knowledge-intensive reasoning tasks: ScienceQA and TabMWP. Chameleon, powered by GPT-4, achieves an 86.54% overall accuracy on ScienceQA, improving the best published few-shot result by 11.37%. On TabMWP, GPT-4-powered Chameleon improves the accuracy by 17.0%, lifting the state of the art to 98.78%. Our analysis also shows that the GPT-4-powered planner exhibits more consistent and rational tool selection via inferring potential constraints from instructions, compared to a ChatGPT-powered planner. The project is available at <https://chameleon-llm.github.io>.